

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-149461

(43) 公開日 平成8年(1996)6月7日

(51) Int.Cl.<sup>8</sup>

識別記号

庁内整理番号

F I

技術表示箇所

H 0 4 N 7/24

G 0 1 B 11/24

G 0 6 T 9/00

K

H 0 4 N 7/ 13

Z

G 0 6 F 15/ 66

3 3 0 H

審査請求 未請求 請求項の数9 O L (全 11 頁)

(21) 出願番号 特願平6-285658

(22) 出願日 平成6年(1994)11月18日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 藤田 雅博

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72) 発明者 田中 勝之

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72) 発明者 佐藤 仁

東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(74) 代理人 弁理士 佐藤 隆久

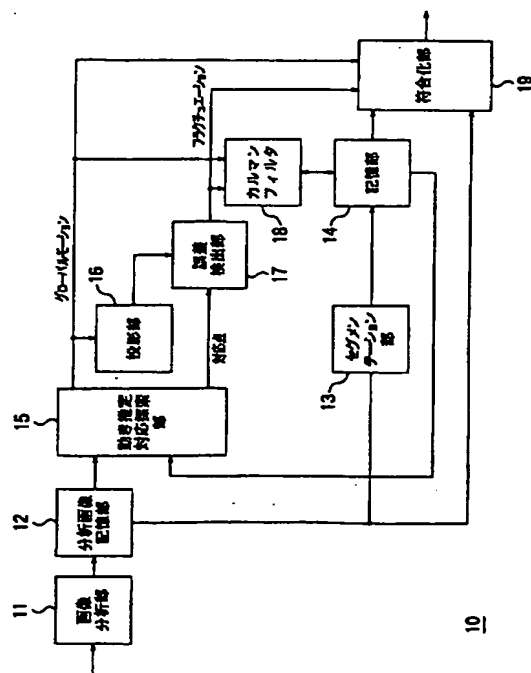
最終頁に続く

(54) 【発明の名称】 動画像処理装置とその方法

(57) 【要約】

【目的】 一連の動画像を構成する各セグメントの3次元形状情報を抽出し、その情報を用いてその一連の動画像を高い圧縮率で符号化する。

【構成】 動画像を画像分析部11で分析し特徴点を求め分析画像記憶部12に記憶する。その1フレーム目よりセグメンテーション部13でセグメントを抽出し記憶部14に記憶する。動き推定・対応探索部15で各セグメントの各フレーム間での動きを推定し、投影部16で各セグメントを移動させ2次元画面上に投影し、誤差検出部17で投影された各特徴点と実際の特徴点の差を求める。その動き推定値と差からカルマンフィルタ18で記憶部14の3次元形状情報を更新する。最初に3次元形状情報と各特徴点の分析値を符号化し、各フレームについてはその3次元形状情報に基づいた動き推定値と各特徴点の差を符号化する。



3

状モデルを抽出する動画像処理装置、および、その形状モデルを用いて高圧縮率で動画像を符号化および復号化する動画像符号化装置、動画像復号化装置、さらに、その符号化された動画像を記録媒体上に記録する動画像記録装置、および、それを再生する動画像再生装置に関する。

#### 【0002】

【従来の技術】動画像系列の中の各物体の3次元モデルを使って動画像を符号化することで、動画像系列を圧縮する方法が提案されている。各物体の3次元形状とその動きがわかれば、元の動画像系列と全く同じ動画像系列が生成できる。そこで、たとえば画像通信において、送信側と受信側で3次元モデルを共有し、送信側で入力画像の動きの情報を検出し、受信側でその動きの情報から画像合成を行えば、画像が再生できる。この場合、動きの情報のみを伝送すればよいことから超低レートでの画像通信が期待できる。具体的には、顔の3次元構造モデルをワイヤフレームに変形し送信側と受信側で共有し、表情などの特徴のみを伝送して顔画像の合成を行う方法が盛んに試みられている。

【0003】しかし、この符号化方法を自然画像に応用する場合には、予め3次元モデルを用意しておくことは不可能であり、与えられた動画像系列から3次元モデルを抽出する必要がある。そのような、動画像系列の中から3次元形状モデルを抽出し、そのモデルを利用して動画像を符号化する方法としては、Hans George Musmann、Michael Hotter、Jorn Ostermannらによる「OBJECT-ORIENTED analysis coding of moving images.」[Signal processing:Image Communication 1(1989):117-138, Elsevier SCIENCE PUBLISHERS B.V.]に開示されている方法がある。この方法によれば、エッジ部分については動きベクトルを求めそれを利用して奥行きを求め、エッジ以外の部分については奥行きを補間して、3次元形状を推測する。また、輝度情報は、3次元上に属性としてマッピングし、補間された3次元形状のデータと共に送出を行っている。

#### 【0004】

【発明が解決しようとする課題】しかし、前述したような方法で3次元形状を推測し、動画像系列を圧縮する方法においては、輝度が急変するエッジ部分さえ動きベクトルを求めることは難しく、動きベクトルから正確な奥行きを推測することは非常に難しかった。したがって、エッジ部分の奥行きを補間して3次元形状を推測しても、正確な3次元形状が求められなかった。また、そのため、実際の形状とのずれがあるため余分な情報が増加して、圧縮率が上げられなかったさらに、前述したような方法では、3次元上に属性としてマッピングした輝度情報を初期情報として伝送するので初期情報量が多いという問題があった。

【0005】したがって、本発明の目的は、動画像系列

4

から忠実度の高い3次元形状モデルを抽出する動画像処理装置を提供することにある。また、その3次元形状モデルを使って高い圧縮率で動画像の符号化が可能な動画像符号化装置、および、それを復号する動画像復号化装置を提供することにある。さらに、その符号化された動画像を記録媒体上に記録する動画像記録装置、および、それを再生する動画像再生装置を提供することにある。

#### 【0006】

【課題を解決するための手段】本発明の動画像処理装置においては、3次元物体の表面データではなく、2次元動画像にしたときに視覚的に重要な点の3次元位置、および、その点の分析値を初期値として持ち、物体の移動を検出しながら、3次元位置情報を獲得するようにした。

【0007】したがって、本発明の動画像処理装置は、入力された連続的な動画像に関する所定の静止画像を分析し、該画像を構成する各セグメントの3次元形状情報を得るモデリング手段と、前記モデリング手段により得られた前記各セグメントの3次元形状情報を記憶する記憶手段と、前記連続的な動画像の各フレーム間における該動画像を構成する各セグメントの3次元動的な動きを、前記記憶手段に記憶されている前記各セグメントの3次元形状情報に基づいて推定する動き推定手段と、前記推定された動きにより各セグメントを3次元的に移動させた結果の位置と実際の位置の差に基づいて、前記記憶手段に記憶されている各セグメントの3次元形状情報を更新する更新手段とを有する。

【0008】好適には、前記モデリング手段は、前記入力された連続的な動画像の各フレームを分析して特徴点を抽出する画像分析手段と、前記連続的な動画像の1フレーム目を構成するセグメントを抽出するセグメンテーション手段と、前記各セグメントごとの前記各特徴点に所定の奥行き値を付加し、前記各セグメントの3次元形状情報の初期値を生成する3次元情報生成手段とを有し、前記動き推定手段は、前記連続的な動画像の各フレーム間における前記各セグメントの3次元動的な動きを、前記記憶手段に記憶されている各セグメントの3次元形状情報に基づいて推定し、前記更新手段は、前記推定された動きにより各セグメントを3次元的に移動させた結果の各特徴点の位置と、実際の各特徴点との位置の差に基づいて、前記記憶手段に記憶されている各セグメントの3次元形状情報を各フレームごとに更新する。

【0009】また好適には、前記画像分析手段は、それぞれ異なる解像度スケールを持つ複数のフィルターで分析する。

【0010】また好適には、前記更新手段は、前記推定された動きにより各セグメントを3次元的に移動させた結果の各特徴点の位置を状態量、当該特徴点の実際の位置を観測量、対応する各特徴点間の位置の差をノイズとし、カルマンフィルタにより前記状態量の最小自乗推定

7

11より入力された1フレーム目の特徴画像データの特徴点の位置と分析値より、セグメンテーションを行い、この画像データを構成しているセグメントを抽出し、各セグメント毎の特徴点の情報を記憶部14に記憶する。このセグメンテーションは、カラー画像から赤・緑・青・明度・色相・彩度の信号、および、テレビ信号に対応したY信号、I信号、Q信号の合計9種類の特徴を抽出し、その特徴に関するヒストグラムに基づいてセグメンテーションを行う再帰的いきい値処理により行う。

【0022】記憶部14は、各セグメントの各特徴点に\*10

$$\{F_{sun} = (p_{sun}, g_{sun}, v_{sun}, a_{sun})\} \quad \dots (1)$$

但し、 $p_{sun}$  は位置情報で、 $p_{sun} = (X_{sun}, Y_{sun}, Z_{sun})$ 、

$s$  はセグメントの番号で、 $s = 1 \sim S$ 、

$u$  はセグメントの番号で、 $u = 1 \sim U_s$ 、

$n$  は各セグメントを構成する点の番号で、 $n = 1 \sim N_{su}$

である。

【0024】なお、確率共分散行列  $v_{sun}$  ( $n = 1 \sim N_{su}$ ,  $u = 1 \sim U_s$ ,  $s = 1 \sim S$ ) は、各エッジを構成する点のちらばりであるので、同一のエッジを構成する各特徴点については同一の値が付される。

【0025】記憶部14に記憶されている情報は、まず、1フレーム目についての情報がセグメンテーション部13より入力され、初期データが生成される。その後、2フレーム目以降の画像データが入力されるごとに、後述するカルマンフィルタ18によりその内容が更新される。

【0026】動き推定・対応探索部15は、前フレームの画像のエッジ画像の各点の情報  $\{F_{sun}\}$  と現フレームのエッジ位置と分析値から、セグメントの動いた量を推定し、前フレームの各点の情報  $\{F_{sun}\}$  と、現フレームのエッジ画像の特徴点の対応付けを行う。その方法について具体的に以下に説明する。まず、図2において ※

$$x_q = f \cdot X_p / Z_p$$

$$y_q = f \cdot Y_p / Z_p$$

【0029】あるセグメント  $s$  ( $s = 1 \sim S$ ) が  $U_s$  個 ( $s = 1 \sim S$ ) のエッジより構成され、その各エッジが  $N_{su}$  個 ( $u = 1 \sim U_s$ ,  $s = 1 \sim S$ ) の点の情報で表され、それら各点の位置は  $p_{sun} = (X_{sun}, Y_{sun}, Z_{sun})$  ( $n = 1 \sim N_{su}$ ) で表されるとする。このセグメントが、相対的にX軸周りに  $\Delta \omega_x$ 、Y軸周りに  $\Delta \omega_y$ 、Z軸周りに  $\Delta \omega_z$  回転し、また、 $\Delta t = (\Delta t_x,$

8

\*ついて、位置情報  $X, Y, Z$  と、分析値  $g$ 、確率共分散行列  $v$ 、付加情報  $a$  を記憶する記憶手段であり、メモリにより構成される。記憶部14に記憶されている情報は、入力された連続シーケンスが  $S$  個のセグメントを有し、各セグメントが  $U_s$  個 ( $s = 1 \sim S$ ) のエッジより構成され、その各エッジが  $N_{su}$  個 ( $u = 1 \sim U_s$ ,  $s = 1 \sim S$ ) の特徴点より構成される場合、式1のように表される。

【0023】

【数1】

\*て、座標系  $XYZ$  はカメラ座標系で、座標系の原点はレンズの中心で、光軸は奥行き方向となる  $Z$  軸と一致させているものとする。このような座標系においては、点  $P$  の像は  $XY$  平面に平行で原点からカメラの焦点距離  $f$  だけ離れた所に設置された平面に投影されたと考えることができる。この投影面上の点  $P$  の像の位置がカメラより入力された画像上の画素の位置となる。その投影面に対して、その面の  $Z$  軸との交点を原点とし、 $X$  軸および  $Y$  軸と平行な座標系  $xy$  を設定する。

30 【0027】 $XYZ$  空間内の点  $P$  の座標を  $p = (X_p, Y_p, Z_p)$ 、点  $P$  の  $xy$  平面上の像である点  $Q$  の座標を  $q = (x_q, y_q)$  とすると、点  $Q$  の座標  $q$  は式2のように表される。

【0028】

【数2】

... (2)

$\Delta t_y, \Delta t_z$ ) だけ平行移動した場合、このセグメントを構成する各点  $p_{sun}$  の移動量  $\Delta p_{sun} = (\Delta X_{sun}, \Delta Y_{sun}, \Delta Z_{sun})$  は、前記各軸周りの回転  $\Delta \omega_x, \Delta \omega_y, \Delta \omega_z$ 、および、平行移動量  $\Delta t$  が小さいとすると、式3のようになる。

【0030】

【数3】

11

$$C_{k+1} = C_k + \Delta C^k$$

$$\Delta p_{sun} = \Sigma \Delta p_{sunk}$$

但し、 $k$  は繰り返し回数である。

【0041】この計算を、 $\Sigma | \Delta q_{sun} |^2$  が予め定めた所定値  $\epsilon$  以下になるまで繰り返すことにより、元の画像の3次元位置情報  $p_{sun}$  に対する新たな画像の対応点  $q_{sun}$  が求められる。

【0042】以上述べたような動き推定・対応探索の方法によれば、物体を剛体と仮定し、回転および平行移動についての6個のパラメータで3次元位置情報を構成する点を拘束することで、個々の点それぞれ独立ではなく、包括的に動き推定・対応探索が行われている。したがって、全体として矛盾のない対応関係が全ての点について得られ、誤対応による3次元位置情報におけるノイズが減る。

【0043】投影部16は、動き推定・対応探索部15により得られた、各セグメントの平行移動量  $t$ 、回転移動量  $\omega$  によって各セグメントの情報  $\{F_{sun}\}$  の位置情報を、3次元空間において平行および回転移動させ、さらに、各セグメントの各点の3次元位置  $p_{sun} = (X_{sun}, Y_{sun}, Z_{sun})$  を画像上の位置  $q_{sun} = (x_{sun}, y_{sun})$  に変換し、得られた画像上の点  $q_{sun}$  に分析値  $g_{sun}$  を与える。3次元位置  $p_{sun}$  から投影点  $q_{sun}$  への変換は式3により行う。

【0044】誤差検出部17は、投影後の各特徴点の情報  $\{F_{sun}\}$  の画像上での位置  $q_{sun}'$  と対応する入力画像のエッジ位置  $q_{sun}$  との差を求め、さらに、その差を量子化し、フラクチュエーションを求める。量子化方法としては、一定の適切な量子化ステップ（たとえば1画素幅）による線形な量子化、いくつかの線形でない量子化ステップを設定した非線形量子化、量子化ステップを固定せず、入力される画像の性質により量子化ステップを適宜変える量子化などがあり、要求される伝送レート、画質に応じて、適切な量子化方法を用いれば良い。たとえば、高圧縮率が要求される場合には、量子化ステップを大きくしたり、画像に直線が多く量子化ノイズによる直線の不連続性が目立つ場合は、非線形量子化を行い、フラクチュエーションの小さい部分の量子化を細かくするようにする。求められたフラクチュエーションは、カルマンフィルタ18に入力される。

【0045】カルマンフィルタ18は、前の画像における各セグメントの各点の情報  $\{F_{sun}\}$  の3次元位置  $p_{sun}$  とそれに対応する入力画像のエッジ位置  $q_{sun}$  から3次元位置  $p_{sun}$  を更新する。カルマンフィルタはノイズを含むシステムにおいて時系列の観測量から状態量の最小自乗推定値を逐次得ることのできるフィルタである。本実施例において、状態量は3次元位置  $p_{sun}$ 、観

12

... (8)

測量である2次元位置  $q_{sun}$  である。2次元位置  $q_{sun}$  には  $\Delta q_{sun}$  の量子化によるノイズが含まれる。また、動き推定値にもノイズが含まれる。初期値の平面上の3次元形状  $|p_{sun}|$  は、カルマンフィルタによりセグメントに動きがあるごとに、実際の3次元形状に近づくように更新されていく。各点の情報  $\{F_{sun}\}$  における確率共分散行列  $v_{sun}$  は  $p_{sun}$  の確率共分散行列 ( $3 \times 3$ ) で  $p_{sun}$  を更新するのに用いられ、同時に確率共分散行列  $v_{sun}$  も更新される。

【0046】以上の、カルマンフィルタにおける更新を、連続的な動画の全フレームについて行くと、最終的に記憶部14には、各セグメントごとの忠実度の高い3次元形状モデルが記憶される。

【0047】次に、前記第1のステップにおいて抽出された3次元形状モデルを用いて、この連続的な動画を符号化する第2のステップについて説明する。第2のステップにおいても、各部の動作は第1のステップと同じである。しかし、第2のステップにおいては、記憶部14に記憶されている最終的な各セグメントの3次元形状情報を用いて動き推定・対応探索を行い、セグメントごとの動きを抽出し、各特徴点の実際の位置との差を求める。

【0048】したがって、まず、動き推定・対応探索部15において、記憶部14に記憶されている各セグメントの3次元形状情報を用いて、分析画像記憶部12に記憶されている各フレームごとの特徴点の位置より、各セグメントの全体の動きと、各特徴点の対応を求める。その求め方は、前記第1のステップの場合と同一である。ここで求められた動きは投影部16および符号化部19に出力される。

【0049】そして、投影部16は、動き推定・対応探索部15により得られた各セグメントの移動量によって各セグメントの位置情報を、3次元空間において平行および回転移動させ、さらに、各セグメントの各点の3次元位置を、画像上の位置に変換する。誤差検出部17は、投影後の各特徴点の画像上での位置と対応する入力画像のエッジ位置との差を求め、さらに、その差を量子化し、フラクチュエーションを求める。求められたフラクチュエーションは、符号化部19に出力される。

【0050】符号化部19は、入力された動画画像系列の情報を符号化し、伝送路に送出する。符号化部19は、各連続画像シーケンスについて、まず、記憶部14に記憶されている3次元形状情報、および、分析値、を符号化する。また、各フレームの画像データについては、各

15

16

【図3】図1に示した動画像符号化装置の動き推定・対応探索の方法を説明する図である。

【図4】本発明の一実施例の動画像復号化装置の構成を示すブロック図である。

【符号の説明】

10…動画像符号化装置

11…画像分析部

13…セグメンテーション部

12…分析画像記憶部

14…記憶部

\* 15…動き推定・対応探索部

17…誤差検出部

30…動画像復号化装置

31…復号部

33…動き処理部

35…変形部

\* 37…合成画像記憶部

16…投影部

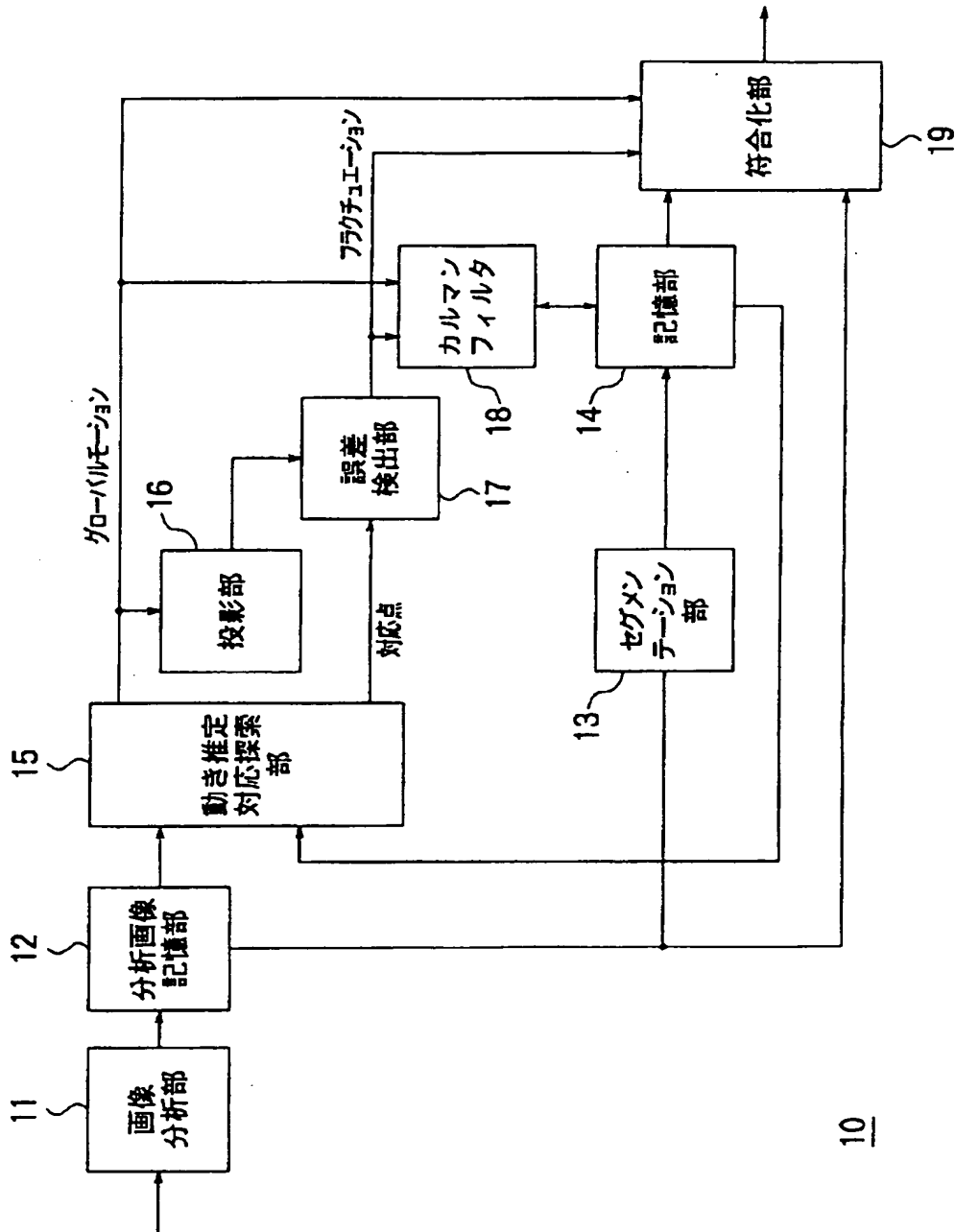
18…カルマンフィルタ

32…記憶部

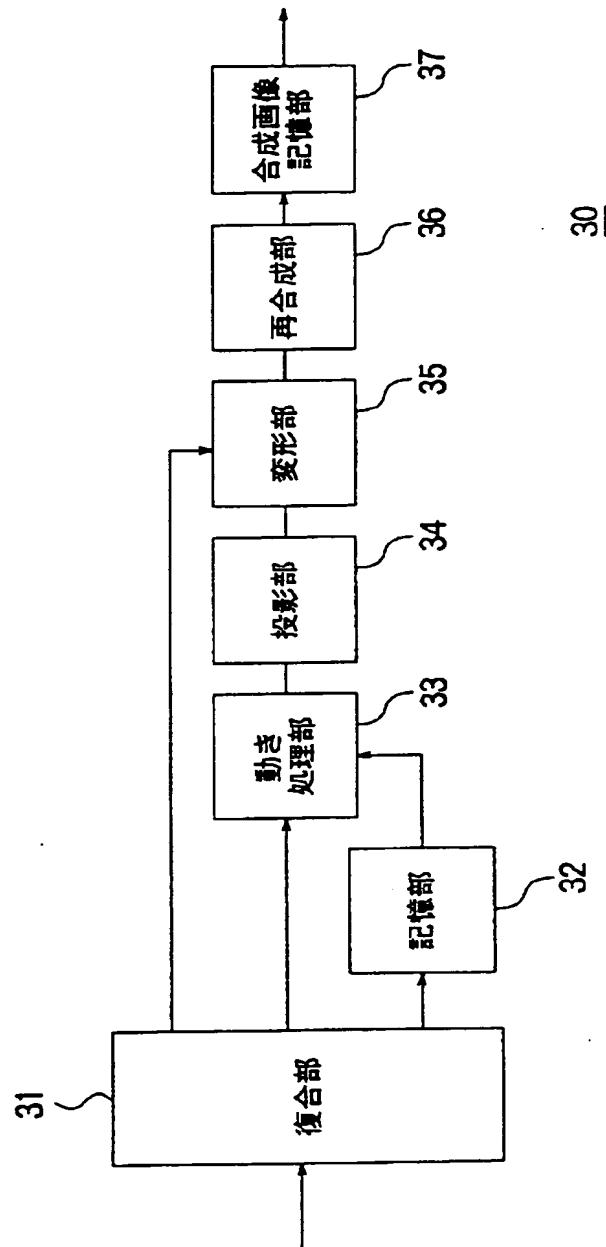
34…投影部

36…再合成部

【図1】



【図 4】



フロントページの続き

(72)発明者 有沢 繁  
東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内